#3

TRANSLATOR'S DECLARATION

I, Paul David Churchill Clarke, B.A., MITI., translator to Messrs. Taylor and Meyer of 20 Kingsmead Road, London, SW2 3JD, Great Britain, verify that I know well both the German and the English language, that I have prepared the attached English translation of 36 pages of the German Patent Application No 199 51 975.7 in the German language with the title:

Neue für das poxB-Gen codierende Nukleotidsequenzen

identified by the code number 990159 BT at the upper left of each page and that the attached English translation of this document is a true and correct translation of the document attached thereto to the best of my knowledge and belief.

I further declare that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements are made with the knowledge that wilful false statements and the like are punishable by fine or imprisonment, or both, under 18 USC 1001, and that such false statements may jeopardize the validity of this document.

Date: 7 January 2000

By: P.D.C. Clarke.

Novel nucleotide sequences coding for the poxB gene

BACKGROUND of the INVENTION

The present invention provides nucleotide sequences from coryneform bacteria coding for the poxB gene and a process for the fermentative production of amino acids, in particular L-lysine, by attenuation of the poxB gene.

~~Prior art~~ 2. Background Information

L-amino acids, in particular lysine, are used in human medicine and in the pharmaceuticals industry, in the food industry and very particularly in animal nutrition.

It is known that amino acids are produced by fermentation of strains of coryneform bacteria, in particular Corynebacterium glutamicum. Due to their great significance, efforts are constantly being made to improve the production process. Improvements to the process may relate to measures concerning fermentation technology, for example stirring and oxygen supply, or to the composition of the nutrient media, such as for example sugar concentration during fermentation, or to working up of the product by, for example, ion exchange chromatography, or to the intrinsic performance characteristics of the microorganism itself.

The performance characteristics of these microorganisms are improved using methods of mutagenesis, selection and mutant selection. In this manner, strains are obtained which are resistant to antimetabolites or are auxotrophic for regulatorily significant amino acids and produce amino acids.

For some years, the methods of recombinant DNA technology have also been used for strain improvement of strains of Corynebacterium which produce L-amino acid.

SUMMARY OF THE INVENTION

Object of the invention

The inventors set themselves the object of providing novel measures for the improved fermentative production of amino acids, in particular L-lysine.

Description of the invention

L-amino acids, in particular lysine, are used in human medicine and in the pharmaceuticals industry, in the food industry and very particularly in animal nutrition. There is accordingly general interest in providing novel improved process for the production of amino acids, in particular L-lysine.

The present invention provides an isolated polynucleotide containing a polynucleotide sequence selected from the group

a) polynucleotide which is at least 70% identical to a polynucleotide which codes for a polypeptide containing the amino acid sequence of SEQ ID no. 2,

b) polynucleotide which codes for a polypeptide which contains an amino acid sequence which is at least 70% identical to the amino acid sequence of SEQ ID no. 2,

c) polynucleotide which is complementary to the polynucleotides of a) or b) and

d) polynucleotide containing at least 15 successive bases of the polynucleotide sequence of a), b) or c).

The present invention also provides the polynucleotide as claimed in claim 1, wherein it preferably comprises a replicable DNA containing:

(i) the nucleotide sequence shown in SEQ ID no. 1, or

(ii) at least one sequence which matches the sequence(i) within the degeneration range of the genetic code, or

(iii) at least one sequence which hybridises with the complementary sequence to sequence (i) or (ii) and optionally

(iv) functionally neutral sense mutations in (i).

The present invention also provides

a polynucleotide according to claim 2, containing the nucleotide sequence as shown in SEQ ID no. 1,

a polynucleotide as claimed in claim 2 which codes for a polypeptide which contains the amino acid sequence as shown in SEQ ID no. 2,

a vector containing the polynucleotide as claimed in claim 1, point d, in particular pCR2.1poxBint, deposited in E. coli DSM 13114

and coryneform bacteria acting as host cell which contain an insertion or deletion in the pox gene.

The present invention also provides polynucleotides which substantially consist of a polynucleotide sequence, which are obtainable by screening by means of hybridisation of a suitable gene library, which contains the complete gene having the polynucleotide sequence according to SEQ ID no. 1, with a probe which contains the sequence of the stated polynucleotide according to SEQ ID no. 1 or a fragment thereof and isolation of the stated DNA sequence.

Polynucleotide sequences according to the invention are suitable as hybridisation probes for RNA, cDNA and DNA in order to isolate full length cDNA which code for the lrp protein and to isolate such cDNA or genes, the sequence of which exhibits a high level of similarity with that of the pyruvate oxidase gene.

Polynucleotide sequences according to the invention are furthermore suitable as primers for the production of DNA of genes which code for pyruvate oxidase by the polymerase chain reaction (PCR).

Such oligonucleotides acting as probes or primers contain at least 30, preferably at least 20, very particularly preferably at least 15 successive nucleotides. Oligonucleotides having a length of at least 40 or 50 bases are also suitable.

"Isolated" means separated from its natural surroundings.

"Polynucleotide" generally denotes polyribonucleotides and polydeoxyribonucleotides, wherein the RNA or DNA may be unmodified or modified.

"Polypeptides" is taken to mean peptides or proteins which contain two or more amino acids joined via peptide bonds.

The polypeptides according to the invention include a polypeptide according to SEQ ID no. 2, in particular those having the biological activity of pyruvate oxidase and also those which are at least 70%, preferably at least 80% and in particular 90% to 95% identical to the polypeptide according to SEQ ID no. 2 and exhibit the stated activity.

The invention furthermore relates to a process for the fermentative production of amino acids, in particular lysine, using coryneform bacteria, which in particular already produce the amino acids, in particular L-lysine, and in which the nucleotide sequences which code for the poxB gene are attenuated, in particular are expressed at a low level.

In this connection, the term "attenuation" means reducing or suppressing the intracellular activity of one or more enzymes (proteins) in a microorganism, which enzymes are coded by the corresponding DNA, for example by using a weak promoter or a gene or allele which codes for a corresponding enzyme which has a low activity or inactivates the corresponding gene or enzyme (protein) and optionally by combining these measures.

The microorganisms, provided by the present invention, may produce amino acids, in particular lysine, from glucose, sucrose, lactose, fructose, maltose, molasses, starch, cellulose or from glycerol and ethanol. The microorganisms may comprise representatives of the coryneform bacteria in particular of the genus Corynebacterium. Within the genus Corynebacterium, Corynebacterium glutamicum may in particular be mentioned, which is known in specialist circles for its ability to produce L-amino acids.

Suitable strains of the genus Corynebacterium, in particular of the species Corynebacterium glutamicum, are in particular the known wild type strains

Corynebacterium glutamicum ATCC13032
Corynebacterium acetoglutamicum ATCC15806
Corynebacterium acetoacidophilum ATCC13870
Corynebacterium melassecola ATCC17965
Corynebacterium thermoaminogenes FERM BP-1539
Brevibacterium flavum ATCC14067
Brevibacterium lactofermentum ATCC13869 and
Brevibacterium divaricatum ATCC14020

and amino acid producing mutants or strains produced therefrom, such as for example

such as for example the L-lysine producing strains

Corynebacterium glutamicum FERM-P 1709
Brevibacterium flavum FERM-P 1708
Brevibacterium lactofermentum FERM-P 1712
Corynebacterium glutamicum FERM-P 6463
Corynebacterium glutamicum FERM-P 6464 and
Corynebacterium glutamicum DSM5714

The inventors succeeded in isolating the novel poxB gene, which codes for the enzyme pyruvate oxidase (EC 1.2.2.2), from C. glutamicum.

The poxB gene or also other genes are isolated from C. glutamicum by initially constructing a gene library of this microorganism in E. coli. The construction of gene libraries is described in generally known textbooks and manuals. Examples which may be mentioned are the textbook by Winnacker, Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Germany, 1990) or the manual by Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989). One very well known gene library is that of E. coli K-12 strain W3110, which was constructed by Kohara et al. (Cell 50, 495-508 (1987)) in λ-vectors. Bathe et al. (Molecular and General Genetics, 252:255-265, 1996) describe a gene library of C. glutamicum ATCC13032, which was constructed using the cosmid vector SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84:2160-2164) in E. coli K-12 strain NM554 (Raleigh et al., 1988, Nucleic Acids Research 16:1563-1575). Börmann et al. (Molecular Microbiology 6(3), 317-326, 1992) also describe a gene library of C. glutamicum ATCC13032, using cosmid pHC79 (Hohn and Collins, Gene 11, 291-298 (1980)). O'Donohue (The Cloning and Molecular Analysis of Four Common Aromatic Amino Acid Biosynthetic Genes from Corynebacterium glutamicum. Ph.D. Thesis, National University of Ireland, Galway, 1997) describes the cloning of C. glutamicum genes using the λ Zap Expression system described by Short et al. (Nucleic Acids Research, 16: 7583).

A gene library of C. glutamicum in E. coli may also be produced using plasmids such as pBR322 (Bolivar, Life Sciences, 25, 807-818 (1979)) or pUC9 (Vieira et al., 1982, Gene, 19:259-268). Suitable hosts are in particular those E. coli strains with restriction and recombination defects, such as for example strain DH5α (Jeffrey H. Miller: "A Short Course in Bacterial Genetics, A Laboratory Manual and

Handbook for Escherichia coli and Related Bacteria", Cold Spring Harbor Laboratory Press, 1992).

The long DNA fragments cloned with the assistance of cosmids or other λ vectors may then in turn be sub-cloned in conventional vectors suitable for DNA sequencing.

DNA sequencing methods are described, inter alia, in Sanger et al. (Proceedings of the National Academy of Sciences of the United States of America USA, 74:5463-5467, 1977).

The resultant DNA sequences may then be investigated using known algorithms or sequence analysis programs, for example Staden's program (Nucleic Acids Research 14, 217-232(1986)), Butler's GCG program (Methods of Biochemical Analysis 39, 74-97 (1998)), Pearson & Lipman's FASTA algorithm (Proceedings of the National Academy of Sciences USA 85,2444-2448 (1988)) or Altschul et al.'s BLAST algorithm (Nature Genetics 6, 119-129 (1994)) and compared with the sequence entries available in publicly accessible databases. Publicly accessible nucleotide sequence databases are, for example, the European Molecular Biologies Laboratories ~~(sic)~~ Biology Laboratory database (EMBL, Heidelberg, Germany) or the National Center for Biotechnology Information database (NCBI, Bethesda, MD, USA).

The novel DNA sequence from C. glutamicum which codes for the poxB gene and, as SEQ ID no. 1, is provided by the present invention, was obtained in this manner. The amino acid sequence of the corresponding protein was furthermore deduced from the above DNA sequence using the methods described above. The resultant amino acid sequence of the poxB gene product is shown in SEQ ID no. 2.

Coding DNA sequences arising from SEQ ID no. 1 due to the degeneracy of the genetic code are also provided by the present invention. DNA sequences which hybridise with SEQ ID no. 1 or parts of SEQ ID no. 1 are similarly provided by

the invention. Finally, DNA sequences produced by the polymerase chain reaction (PCR) using primers obtained from SEQ ID no. 1 are also provided by the present invention.

The person skilled in the art may find instructions for identifying DNA sequences by means of hybridisation inter alia in the manual "The DIG System Users Guide for Filter Hybridization" from Boehringer Mannheim GmbH (Mannheim, Germany, 1993) and in Liebl et al. (International Journal of Systematic Bacteriology (1991) 41: 255-260). The person skilled in the art will find instructions for amplifying DNA sequences by means of the polymerase chain reaction (PCR) inter alia in the textbook by Gait, Oligonucleotide synthesis: a practical approach (IRL Press, Oxford, UK, 1984) and in Newton and Graham, PCR (Spektrum Akademischer Verlag, Heidelberg, Germany, 1994).

The inventors discovered that coryneform bacteria produce L-amino acids, in particular L-lysine, in an improved manner once the poxB has been attenuated.

Attenuation may be achieved by reducing or suppressing either expression of the poxB gene or the catalytic properties of the enzyme protein. These measures may optionally be combined.

Reduced gene expression may be achieved by appropriate control of the culture or by genetic modification (mutation) of the signal structures for gene expression. Signal structures for gene expression are, for example, repressor genes, activator genes, operators, promoters, attenuators, ribosome binding sites, the start codon and terminators. The person skilled in the art will find information in this connection for example in patent application WO 96/15246, in Boyd & Murphy (Journal of Bacteriology 170: 5949 (1988)), in Voskuil & Chambliss (Nucleic Acids Research 26: 3548 (1998)), in Jensen & Hammer (Biotechnology and Bioengineering 58: 191 (1998)),

in Patek et al. (Microbiology 142: 1297 (1996)) and in known textbooks of genetics and molecular biology, such as for example the textbook by Knippers ("Molekulare Genetik", 6th edition, Georg Thieme Verlag, Stuttgart, Germany, 1995) or by Winnacker ("Gene und Klone", VCH Verlagsgesellschaft, Weinheim, Germany, 1990).

Mutations which give rise to a change or reduction in the catalytic properties of enzyme proteins are known from the prior art; examples which may be mentioned are the papers by Qiu and Goodman (Journal of Biological Chemistry 272: 8611-8617 (1997)), Sugimoto et al. (Bioscience Biotechnology and Biochemistry 61: 1760-1762 (1997)) and Möckel ("Die Threonindehydratase aus Corynebacterium glutamicum: Aufhebung der allosterischen Regulation und Struktur des Enzyms", Berichte des Forschungszentrums Jülichs, Jül-2906, ISSN09442952, Jülich, Germany, 1994). Summary explanations may be found in known textbooks of genetics and molecular biology, such as for example that by Hagemann ("Allgemeine Genetik", Gustav Fischer Verlag, Stuttgart, 1986).

Mutations which may be considered are transitions, transversions, insertions and deletions. Depending upon the effect of exchanging the amino acids upon enzyme activity, the mutations are known as missense mutations or nonsense mutations. Insertions or deletions of at least one base pair in a gene give rise to frame shift mutations, as a result of which the incorrect amino acids are inserted or translation terminates prematurely. Deletions of two or more codons typically result in a complete breakdown of enzyme activity. Instructions for producing such mutations belong to the prior art and may be found in known textbooks of genetics and molecular biology, such as for example the textbook by Knippers ("Molekulare Genetik", 6th edition, Georg Thieme Verlag, Stuttgart, Germany, 1995), by Winnacker ("Gene und Klone", VCH Verlagsgesellschaft,

Weinheim, Germany, 1990) or by Hagemann ("Allgemeine Genetik", Gustav Fischer Verlag, Stuttgart, 1986).

One example of a plasmid with the assistance of which insertion mutagenesis of the poxB gene may be performed is pCR2.1poxBint (Figure 1).

Plasmid pCR2.1poxBint consists of the plasmid pCR2.1-TOPO described by Mead et al. (Bio/Technology 9:657-663 (1991)), into which an internal fragment of the poxB gene, shown in SEQ ID no. 3, has been incorporated. After transformation and homologous recombination into the chromosomal poxB gene (insertion), this plasmid results in a total loss of enzyme function. By way of example, the strain DSM5715::pCR2.1poxBint, the pyruvate oxidase of which is switched off, was produced in this manner. Further instructions and explanations relating to insertion mutagenesis may be found, for example, in Schwarzer and Pühler (Bio/Technology 9,84-87 (1991)) or Fitzpatrick et al. (Applied Microbiology and Biotechnology 42, 575-580 (1994)).

It may additionally be advantageous for the production of L-amino acids, in particular L-lysine, in addition to attenuating the poxB gene, to amplify, in particular to overexpress, one or more enzymes of the particular biosynthetic pathway, of glycolysis, of anaplerotic metabolism, of the citric acid cycle or of amino acid export.

Thus, for example, for the production of L-lysine

- the dapA gene (EP-B 0 197 335) which codes for dihydropicolinate synthase may simultaneously be overexpressed, or
- the dapD gene (Wehrmann et al., Journal of Bacteriology 180, 3159-3165 (1998)) which codes for

tetradihydropicolinate succinylase may simultaneously be overexpressed, or

- the dapE gene (Wehrmann et al., Journal of Bacteriology 177: 5991-5993 (1995)) which codes for succinyldiaminopimelate desuccinylase may simultaneously be overexpressed, or

- the gap gene (Eikmanns (1992), Journal of Bacteriology 174:6076-6086) which codes for glyceraldehyde 3-phosphate dehydrogenase may simultaneously be overexpressed, or

- the pyc gene (Eikmanns (1992), Journal of Bacteriology 174:6076-6086) which codes for pyruvate carboxylase may simultaneously be overexpressed, or

- the mqo gene (Molenaar et al., European Journal of Biochemistry 254, 395 - 403 (1998)) which codes for malate:quinone oxidoreductase may simultaneously be overexpressed, or

- the lysE gene (DE-A-195 48 222) which codes for lysine export may simultaneously be overexpressed.

It may furthermore be advantageous for the production of amino acids, in particular L-lysine, in addition to attenuating the poxB gene, to suppress unwanted secondary reactions (Nakayama: "Breeding of Amino Acid Producing Micro-organisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).

The microorganisms containing the polynucleotide according to claim 1 are also provided by the invention and may be cultured continuously or discontinuously using the batch process or the fed batch process or repeated fed batch process for the purpose of producing L-amino acids, in particular L-lysine. A summary of known culture methods is given in the textbook by Chmiel (Bioprozesstechnik 1.

Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) or in the textbook by Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)).

The culture medium to be used must adequately satisfy the requirements of the particular strains. Culture media for various microorganisms are described in "Manual of Methods for General Bacteriology" from the American Society for Bacteriology (Washington D.C., USA, 1981). Carbon sources which may be used include sugars and carbohydrates, such as for example glucose, sucrose, lactose, fructose, maltose, molasses, starch and cellulose, oils and fats, such as for example soya oil, sunflower oil, peanut oil and coconut oil, fatty acids, such as for example palmitic acid, stearic acid and linoleic acid, alcohols, such as for example glycerol and ethanol, and organic acids, such as for example acetic acid. These substances may be used individually or as a mixture. Nitrogen sources which may be used comprise organic compounds containing nitrogen, such as peptones, yeast extract, meat extract, malt extract, corn steep liquor, soya flour and urea or inorganic compounds, such as ammonium sulfate, ammonium chloride, ammonium phosphate, ammonium carbonate and ammonium nitrate. The nitrogen sources may be used individually or as a mixture. Phosphorus sources which may be used are phosphoric acid, potassium dihydrogen phosphate or dipotassium hydrogen phosphate or the corresponding salts containing sodium. The culture medium must furthermore contain metal salts, such as for example magnesium sulfate or iron sulfate, which are necessary for growth. Finally, essential growth-promoting substances such as amino acids and vitamins may also be used in addition to the above-stated substances. Suitable precursors may furthermore be added to the culture medium. The stated materials may be added to the culture in the form of a single batch or may be supplied in a suitable manner during culturing.

Basic compounds, such as sodium hydroxide, potassium hydroxide, ammonia or ammonia water, or acidic compounds, such as phosphoric acid or sulfuric acid, are used appropriately to control the pH of the culture. Antifoaming agents, such as for example fatty acid polyglycol esters, may be used to control foaming. Suitable selectively acting substances, such as for example antibiotics, may be added to the medium in order to maintain plasmid stability. Oxygen or gas mixtures containing oxygen, such as for example air, are introduced into the culture in order to maintain aerobic conditions. The temperature of the culture is normally from 20°C to 45°C and preferably from 25°C to 40°C. The culture is continued until the maximum quantity of the desired amino acid has formed. This objective is normally achieved within 10 hours to 160 hours.

Methods for determining L-amino acids are known from the prior art. Analysis may proceed by anion exchange chromatography with subsequent ninhydrin derivatisation, as described in Spackman et al. (Analytical Chemistry, 30, (1958), 1190) or by reversed phase HPLC, as described in Lindroth et al. (Analytical Chemistry (1979) 51: 1167-1174).

The following microorganism has been deposited with Deutschen Sammlung für Mikrorganismen und Zellkulturen (DSMZ, Braunschweig, Germany) in accordance with the Budapest Treaty:

- Escherichia coli strain DH5α/pCR2.1poxBint as DSM 13114.

Examples DETAILED DESCRIPTION OF THE INVENTION

The present invention is illustrated in greater detail by the following practical examples.

Example 1

Production of a genomic cosmid gene library from Corynebacterium glutamicum ATCC13032

Chromosomal DNA from Corynebacterium glutamicum ATCC13032 was isolated as described in Tauch et al., (1995, Plasmid 33:168-179) and partially cleaved with the restriction enzyme Sau3AI (Amersham Pharmacia, Freiburg, Germany, product description Sau3AI, code no. 27-0913-02). The DNA fragments were dephosphorylated with shrimp alkaline phosphatase (Roche Molecular Biochemicals, Mannheim, Germany, product description SAP, code no. 1758250). The DNA of cosmid vector SuperCos1 (Wahl et al. (1987) Proceedings of the National Academy of Sciences USA 84:2160-2164), purchased from Stratagene (La Jolla, USA, product description SuperCos1 Cosmid Vector Kit, code no. 251301) was cleaved with the restriction enzyme XbaI (Amersham Pharmacia, Freiburg, Germany, product description XbaI, Code no. 27-0948-02) and also dephosphorylated with shrimp alkaline phosphatase. The cosmid DNA was then cleaved with the restriction enzyme BamHI (Amersham Pharmacia, Freiburg, Germany, product description BamHI, code no. 27-0868-04). Cosmid DNA treated in this manner was mixed with the treated ATCC 13032 DNA and the batch was treated with T4 DNA ligase (Amersham Pharmacia, Freiburg, Germany, product description T4 DNA Ligase, code no. 27-0870-04). The ligation mixture was then packed in phages using Gigapack II XL Packing Extracts (Stratagene, La Jolla, USA, product description Gigapack II XL Packing Extract, code no. 200217). E. coli strain NM554 (Raleigh et al. 1988, Nucleic Acid Res. 16:1563-1575) was infected by suspending the cells in 10 mM $MgSO_4$ and mixing them with an

aliquot of the phage suspension. The cosmid library was infected and titred as described in Sambrook et al. (1989, Molecular Cloning: A laboratory Manual, Cold Spring Harbor), wherein the cells were plated out on LB agar (Lennox, 1955, Virology, 1:190) + 100µg/ml of ampicillin. After overnight incubation at 37°C, individual recombinant clones were selected.

Example 2

Isolation and sequencing of the poxB gene

Cosmid DNA from an individual colony was isolated in accordance with the manufacturer's instructions using the Qiaprep Spin Miniprep Kit (Product No. 27106, Qiagen, Hilden, Germany) and partially cleaved with the restriction enzyme Sau3AI (Amersham Pharmacia, Freiburg, Germany, product description Sau3AI, code no. 27-0913-02). The DNA fragments were dephosphorylated with shrimp alkaline phosphatase (Roche Molecular Biochemicals, Mannheim, Germany, product description SAP, code no. 1758250). Once separated by gel electrophoresis, the cosmid fragments of a size of approx. 1500 to 2000 bp were isolated using the QiaExII Gel Extraction Kit (product no. 20021, Qiagen, Hilden, Germany). The DNA of the sequencing vector pZero-1 purchased from Invitrogen (Groningen, Netherlands, product description Zero Background Cloning Kit, product no. K2500-01) was cleaved with the restriction enzyme BamHI (Amersham Pharmacia, Freiburg, Germany, product description BamHI, Product No. 27-0868-04). Ligation of the cosmid fragments into the sequencing vector pZero-1 was performed as described by Sambrook et al. (1989, Molecular Cloning: A laboratory Manual, Cold Spring Harbor), wherein the DNA mixture was incubated overnight with T4 ligase (Pharmacia Biotech, Freiburg, Germany). This ligation mixture was then electroporated into the E. coli strain DH5aMCR (Grant,

1990, Proceedings of the National Academy of Sciences U.S.A., 87:4645-4649) (Tauch et al. 1994, FEMS Microbiol Letters, 123:343-7) and plated out onto LB agar (Lennox, 1955, Virology, 1:190) + 50 µg/ml of Zeocin. Plasmids of the recombinant clones were prepared using the Biorobot 9600 (product no. 900200, Qiagen, Hilden, Germany, Germany). Sequencing was performed using the dideoxy chain termination method according to Sanger et al. (1977, Proceedings of the National Academies of Sciences U.S.A., 74:5463-5467) as modified by Zimmermann et al. (1990, Nucleic Acids Research, 18:1067). The "RR dRhodamin Terminator Cycle Sequencing Kit" from PE Applied Biosystems (product no. 403044, Weiterstadt, Germany) was used. Separation by gel electrophoresis and analysis of the sequencing reaction was performed in a "Rotiphorese NF" acrylamide/bisacrylamide gel (29:1) (product no. A124.1, Roth, Karlsruhe, Germany) using the "ABI Prism 377" sequencer from PE Applied Biosystems (Weiterstadt, Germany).

The resultant raw sequence data were then processing using the Staden software package (1986, Nucleic Acids Research, 14:217-231), version 97-0. The individual sequences of the pZero 1 derivatives were assembled into a cohesive contig. Computer-aided coding range analysis was performed using XNIP software (Staden, 1986, Nucleic Acids Research, 14:217-231). Further analysis was performed using the "BLAST search programs" (Altschul et al., 1997, Nucleic Acids Research, 25:3389-3402), against the non-redundant database of the "National Center for Biotechnology Information" (NCBI, Bethesda, MD, USA).

The resultant nucleotide sequence is stated in SEQ ID no. 1 . Analysis of the nucleotide sequence revealed an open reading frame of 1737 base pairs, which was designated the poxB gene. The poxB gene codes for a polypeptide of 579 amino acids.

Example 3

Production of an integration vector for integration mutagenesis of the poxB gene

Chromosomal DNA was isolated from strain ATCC 13032 using the method of Eikmanns et al. (Microbiology 140: 1817 - 1828 (1994)). On the basis of the sequence of the poxB gene for C. glutamicum known from Example 2, the following oligonucleotides were selected for the polymerase chain reaction:

poxBint1:
5' TGC GAG ATG GTG AAT GGT GG 3'
poxBint2:
5' GCA TGA GGC AAC GCA TTA GC 3'

The stated primers were synthesised by the company MWG Biotech (Ebersberg, Germany) and the PCR reaction performed in accordance with the standard PCR method of Innis et al. (PCR protocols. A guide to methods and applications, 1990, Academic Press) using Pwo polymerase from Boehringer. A DNA fragment of approx. 0.9 kb in size, which bears an internal fragment of the poxB gene and is shown in SEQ ID no. 3, was isolated with the assistance of the polymerase chain reaction.

The amplified DNA fragment was ligated into the vector pCR2.1-TOPO (Mead at al. (1991) Bio/Technology 9:657-663) using the TOPO TA Cloning Kit from Invitrogen Corporation (Carlsbad, CA, USA; catalogue no. K4500-01). The E. coli strain DH5α was then electroporated with the ligation batch (Hanahan, in DNA cloning. A practical approach. Vol.I. IRL-Press, Oxford, Washington DC, USA, 1985). Plasmid-bearing cells were selected by plating the transformation batch out onto LB agar (Sambrook et al., Molecular cloning: a laboratory manual. 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989) which had been

DJS 08-22-01

supplemented with 25 mg/l of kanamycin. Plasmid DNA was isolated from a transformant using the QIAprep Spin Miniprep Kit from Qiagen and verified by restriction with the restriction enzyme EcoRI and subsequent agarose gel electrophoresis (0.8%). The plasmid was named pCR2.1poxBint.

Example 4

Integration mutagenesis of the poxB gene into the lysine producer DSM 5715

The vector named pCR2.1poxBint in Example 2 was electroporated into Corynebacterium glutamicum DSM 5715 using the electroporation method of Tauch et al. (FEMS Microbiological Letters, 123:343-347 (1994)). Strain DSM 5715 is an AEC-resistant lysine producer. The vector pCR2.1poxBint cannot independently replicate in DSM 5715 and is only retained in the cell if it has been integrated into the chromosome of DSM 5715. Clones with pCR2.1poxBint integrated into the chromosome were selected by plating the electroporation batch out onto LB agar (Sambrook et al., Molecular cloning: a laboratory manual. 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.) which had been supplemented with 15 mg/l of kanamycin. Integration was detected by labelling the poxBint fragment with the Dig hybridisation kit from Boehringer using the method according to "The DIG System Users Guide for Filter Hybridization" from Boehringer Mannheim GmbH (Mannheim, Germany, 1993). Chromosomal DNA of a potential integrant was isolated using the method according to Eikmanns et al. (Microbiology 140: 1817 - 1828 (1994)) and cut in each case with the restriction enzymes SalI, SacI and HinDIII. The resultant fragments were separated by agarose gel electrophoresis and hybridised at 68°C using the Dig hybridisation kit from Boehringer. The plasmid named

pCR2.1poxBint in Example 3 had been inserted within the chromosomal poxB gene in the chromosome of DSM 5715. The strain was designated DSM5715::pCR2.1poxBint.

Example 5

Production of lysine

The C. glutamicum strain DSM5715::pCR2.1poxBint obtained in Example 3 was cultured in a nutrient medium suitable for the production of lysine and the lysine content of the culture supernatant was determined.

To this end, the strain was initially incubated for 24 hours at 33°C on an agar plate with the appropriate antibiotic (brain/heart agar with kanamycin (25 mg/l)). Starting from this agar plate culture, a preculture was inoculated (10 ml of medium in a 100 ml Erlenmeyer flask). The complete medium CgIII was used as the medium for this preculture. Kanamycin (25 ml/l) was added to this medium. The preculture was incubated for 48 hours at 33°C on a shaker at 240 rpm. A main culture was inoculated from this preculture, such that the initial optical density (OD, 660 nm) of the main culture was 0.1 OD. Medium MM was used for the main culture.

Medium MM

CSL (Corn Steep Liquor)	5 g/l
MOPS	20 g/l
Glucose (separately autoclaved)	50 g/l
Salts:	
$(NH_4)_2SO_4$)	25 g/l
KH_2PO_4	0.1 g/l
$MgSO_4$ * 7 H_2O	1.0 g/l
$CaCl_2$ * 2 H_2O	10 mg/l
$FeSO_4$ * 7 H_2O	10 mg/l
$MnSO_4$ * H_2O	5.0mg/l
Biotin (sterile-filtered)	0.3 mg/l
Thiamine*HCl (sterile-filtered)	0.2 mg/l
Leucine (sterile-filtered)	0.1 g/l
$CaCO_3$	25 g/l

CSL, MOPS and the salt solution are adjusted to pH 7 with ammonia solution and autoclaved. The sterile substrate and vitamin solutions, together with the dry-autoclaved $CaCO_3$ are then added.

Culturing is performed in a volume of 10 ml in a 100 ml Erlenmeyer flask with flow spoilers. Kanamycin (25 ml/l) was added. Culturing was performed at 33°C and 80% atmospheric humidity.

BRIEF DESCRIPTION OF THE DRAWING

~~After 48 hours, the OD was determined at a measurement~~ wavelength of 660 nm using a Biomek 1000 (Beckmann Instruments GmbH, Munich). The quantity of lysine formed was determined using an amino acid analyser from Eppendorf-BioTronik (Hamburg, Germany) by ion exchange chromatography and post-column derivatisation with ninhydrin detection.

Table 1 shows the result of the test.

Table 1

Strain	OD(660)	Lysine HCl 5 g/l
DSM5715	13.1	9.5
DSM5715::pCR2.1poxBint	12.5	12.9

The following Figures are attached:

Figure 1: Map of the plasmid pCR2.1poxBint.

The abbreviations and terms used have the following meanings.

ColE1 ori:	Replication origin of the plasmid ColE1
lacZ:	5' end of the β-galactosidase gene
f1 ori:	Replication origin of the f1 phage
KmR:	Kanamycin resistance
ApR:	Ampicillin resistance
BamHI:	Restriction site of the restriction enzyme BamHI
EcoRI:	Restriction site of the restriction enzyme EcoRI:
poxBint2:	Internal fragment of the poxB gene

SEQUENCE LISTING

<110> Degussa-Hüls AG

<120>Novel nucleotide sequences coding for the poxB gene

<130> 990159 BT

<140>
<141>

<160> 3

<170> PatentIn Ver. 2,1

<210> 1
<211> 2160
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (327)..(2063)

<220>
<221> -35_signal
<222> (227)..(232)

<220>
<221> -10_signal
<222> (256)..(261)

<400> 1
ttagaggcga ttctgtgagg tcacttttg tggggtcggg gtctaaattt ggccagtttt 60

cgaggcgacc agacaggcgt gcccacgatg tttaaatagg cgatcggtgg gcatctgtgt 120

ttggtttcga cgggctgaaa ccaaaccaga ctgcccagca acgacggaaa tcccaaaagt 180

gggcatccct gtttggtacc gagtacccac ccgggcctga aactccctgg caggcgggcg 240

aagcgtggca acaactggaa tttaagagca caattgaagt cgcaccaagt taggcaacac 300

aatagccata acgttgagga gttcag atg gca cac agc tac gca gaa caa tta 353
Met Ala His Ser Tyr Ala Glu Gln Leu
1 5

att gac act ttg gaa gct caa ggt gtg aag cga att tat ggt ttg gtg 401
Ile Asp Thr Leu Glu Ala Gln Gly Val Lys Arg Ile Tyr Gly Leu Val
10 15 20 25

ggt gac agc ctt aat ccg atc gtg gat gct gtc cgc caa tca gat att 449
Gly Asp Ser Leu Asn Pro Ile Val Asp Ala Val Arg Gln Ser Asp Ile
30 35 40

gag tgg gtg cac gtt cga aat gag gaa gcg gcg gcg ttt gca gcc ggt 497
Glu Trp Val His Val Arg Asn Glu Glu Ala Ala Ala Phe Ala Ala Gly
45 50 55

gcg gaa tcg ttg atc act ggg gag ctg gca gta tgt gct gct tct tgt 545
Ala Glu Ser Leu Ile Thr Gly Glu Leu Ala Val Cys Ala Ala Ser Cys
60 65 70

```
ggt cct gga aac aca cac ctg att cag ggt ctt tat gat tcg cat cga   593
Gly Pro Gly Asn Thr His Leu Ile Gln Gly Leu Tyr Asp Ser His Arg
     75                  80                  85

aat ggt gcg aag gtg ttg gcc atc gct agc cat att ccg agt gcc cag   641
Asn Gly Ala Lys Val Leu Ala Ile Ala Ser His Ile Pro Ser Ala Gln
 90                  95                 100                 105

att ggt tcg acg ttc ttc cag gaa acg cat ccg gag att ttg ttt aag   689
Ile Gly Ser Thr Phe Phe Gln Glu Thr His Pro Glu Ile Leu Phe Lys
                110                 115                 120

gaa tgc tct ggt tac tgc gag atg gtg aat ggt ggt gag cag ggt gaa   737
Glu Cys Ser Gly Tyr Cys Glu Met Val Asn Gly Gly Glu Gln Gly Glu
            125                 130                 135

cgc att ttg cat cac gcg att cag tcc acc atg gcg ggt aaa ggt gtg   785
Arg Ile Leu His His Ala Ile Gln Ser Thr Met Ala Gly Lys Gly Val
        140                 145                 150

tcg gtg gta gtg att cct ggt gat atc gct aag gaa gac gca ggt gac   833
Ser Val Val Val Ile Pro Gly Asp Ile Ala Lys Glu Asp Ala Gly Asp
    155                 160                 165

ggt act tat tcc aat tcc act att tct tct ggc act cct gtg gtg ttc   881
Gly Thr Tyr Ser Asn Ser Thr Ile Ser Ser Gly Thr Pro Val Val Phe
170                 175                 180                 185

ccg gat cct act gag gct gca gcg ctg gtg gag gcg att aac aac gct   929
Pro Asp Pro Thr Glu Ala Ala Ala Leu Val Glu Ala Ile Asn Asn Ala
                190                 195                 200

aag tct gtc act ttg ttc tgc ggt gcg ggc gtg aag aat gct cgc gcg   977
Lys Ser Val Thr Leu Phe Cys Gly Ala Gly Val Lys Asn Ala Arg Ala
            205                 210                 215

cag gtg ttg gag ttg gcg gag aag att aaa tca ccg atc ggg cat gcg   1025
Gln Val Leu Glu Leu Ala Glu Lys Ile Lys Ser Pro Ile Gly His Ala
        220                 225                 230

ctg ggt ggt aag cag tac atc cag cat gag aat ccg ttt gag gtc ggc   1073
Leu Gly Gly Lys Gln Tyr Ile Gln His Glu Asn Pro Phe Glu Val Gly
    235                 240                 245

atg tct ggc ctg ctt ggt tac ggc gcc tgc gtg gat gcg tcc aat gag   1121
Met Ser Gly Leu Leu Gly Tyr Gly Ala Cys Val Asp Ala Ser Asn Glu
250                 255                 260                 265

gcg gat ctg ctg att cta ttg ggt acg gat ttc cct tat tct gat ttc   1169
Ala Asp Leu Leu Ile Leu Leu Gly Thr Asp Phe Pro Tyr Ser Asp Phe
                270                 275                 280

ctt cct aaa gac aac gtt gcc cag gtg gat atc aac ggt gcg cac att   1217
Leu Pro Lys Asp Asn Val Ala Gln Val Asp Ile Asn Gly Ala His Ile
            285                 290                 295

ggt cga cgt acc acg gtg aag tat ccg gtg acc ggt gat gtt gct gca   1265
Gly Arg Arg Thr Thr Val Lys Tyr Pro Val Thr Gly Asp Val Ala Ala
        300                 305                 310

aca atc gaa aat att ttg cct cat gtg aag gaa aaa aca gat cgt tcc   1313
Thr Ile Glu Asn Ile Leu Pro His Val Lys Glu Lys Thr Asp Arg Ser
    315                 320                 325
```

```
ttc ctt gat cgg atg ctc aag gca cac gag cgt aag ttg agc tcg gtg   1361
Phe Leu Asp Arg Met Leu Lys Ala His Glu Arg Lys Leu Ser Ser Val
330                 335                 340                 345

gta gag acg tac aca cat aac gtc gag aag cat gtg cct att cac cct   1409
Val Glu Thr Tyr Thr His Asn Val Glu Lys His Val Pro Ile His Pro
                350                 355                 360

gaa tac gtt gcc tct att ttg aac gag ctg gcg gat aag gat gcg gtg   1457
Glu Tyr Val Ala Ser Ile Leu Asn Glu Leu Ala Asp Lys Asp Ala Val
            365                 370                 375

ttt act gtg gat acc ggc atg tgc aat gtg tgg cat gcg agg tac atc   1505
Phe Thr Val Asp Thr Gly Met Cys Asn Val Trp His Ala Arg Tyr Ile
        380                 385                 390

gag aat ccg gag gga acg cgc gac ttt gtg ggt tca ttc cgc cac ggc   1553
Glu Asn Pro Glu Gly Thr Arg Asp Phe Val Gly Ser Phe Arg His Gly
    395                 400                 405

acg atg gct aat gcg ttg cct cat gcg att ggt gcg caa agt gtt gat   1601
Thr Met Ala Asn Ala Leu Pro His Ala Ile Gly Ala Gln Ser Val Asp
410                 415                 420                 425

cga aac cgc cag gtg atc gcg atg tgt ggc gat ggt ggt ttg ggc atg   1649
Arg Asn Arg Gln Val Ile Ala Met Cys Gly Asp Gly Gly Leu Gly Met
                430                 435                 440

ctg ctg ggt gag ctt ctg acc gtt aag ctg cac caa ctt ccg ctg aag   1697
Leu Leu Gly Glu Leu Leu Thr Val Lys Leu His Gln Leu Pro Leu Lys
            445                 450                 455

gct gtg gtg ttt aac aac agt tct ttg ggc atg gtg aag ttg gag atg   1745
Ala Val Val Phe Asn Asn Ser Ser Leu Gly Met Val Lys Leu Glu Met
        460                 465                 470

ctc gtg gag gga cag cca gaa ttt ggt act gac cat gag gaa gtg aat   1793
Leu Val Glu Gly Gln Pro Glu Phe Gly Thr Asp His Glu Glu Val Asn
    475                 480                 485

ttc gca gag att gcg gcg gct gcg ggt atc aaa tcg gta cgc atc acc   1841
Phe Ala Glu Ile Ala Ala Ala Ala Gly Ile Lys Ser Val Arg Ile Thr
490                 495                 500                 505

gat ccg aag aaa gtt cgc gag cag cta gct gag gca ttg gca tat cct   1889
Asp Pro Lys Lys Val Arg Glu Gln Leu Ala Glu Ala Leu Ala Tyr Pro
                510                 515                 520
```

```
gga cct gta ctg atc gat atc gtc acg gat cct aat gcg ctg tcg atc   1937
Gly Pro Val Leu Ile Asp Ile Val Thr Asp Pro Asn Ala Leu Ser Ile
            525                 530                 535

cca cca acc atc acg tgg gaa cag gtc atg gga ttc agc aag gcg gcc   1985
Pro Pro Thr Ile Thr Trp Glu Gln Val Met Gly Phe Ser Lys Ala Ala
        540                 545                 550

acc cga acc gtc ttt ggt gga gga gta gga gcg atg atc gat ctg gcc   2033
Thr Arg Thr Val Phe Gly Gly Gly Val Gly Ala Met Ile Asp Leu Ala
    555                 560                 565

cgt tcg aac ata agg aat att cct act cca tgatgattga tacacctgct     2083
Arg Ser Asn Ile Arg Asn Ile Pro Thr Pro
570                 575

gttctcattg accgcgagcg cttaactgcc aacatttcca ggatggcagc tcacgccggt 2143

gcccatgaga ttgccct                                                2160
```

<210> 2
<211> 579
<212> PRT
<213> Corynebacterium glutamicum

<400> 2

```
Met Ala His Ser Tyr Ala Glu Gln Leu Ile Asp Thr Leu Glu Ala Gln
  1               5                  10                  15

Gly Val Lys Arg Ile Tyr Gly Leu Val Gly Asp Ser Leu Asn Pro Ile
            20                  25                  30

Val Asp Ala Val Arg Gln Ser Asp Ile Glu Trp Val His Val Arg Asn
        35                  40                  45

Glu Glu Ala Ala Ala Phe Ala Ala Gly Ala Glu Ser Leu Ile Thr Gly
    50                  55                  60

Glu Leu Ala Val Cys Ala Ala Ser Cys Gly Pro Gly Asn Thr His Leu
65                  70                  75                  80

Ile Gln Gly Leu Tyr Asp Ser His Arg Asn Gly Ala Lys Val Leu Ala
                85                  90                  95

Ile Ala Ser His Ile Pro Ser Ala Gln Ile Gly Ser Thr Phe Phe Gln
            100                 105                 110

Glu Thr His Pro Glu Ile Leu Phe Lys Glu Cys Ser Gly Tyr Cys Glu
        115                 120                 125

Met Val Asn Gly Gly Glu Gln Gly Glu Arg Ile Leu His His Ala Ile
    130                 135                 140

Gln Ser Thr Met Ala Gly Lys Gly Val Ser Val Val Val Ile Pro Gly
145                 150                 155                 160
```

```
Asp Ile Ala Lys Glu Asp Ala Gly Asp Gly Thr Tyr Ser Asn Ser Thr
                165                 170                 175

Ile Ser Ser Gly Thr Pro Val Val Phe Pro Asp Pro Thr Glu Ala Ala
            180                 185                 190

Ala Leu Val Glu Ala Ile Asn Asn Ala Lys Ser Val Thr Leu Phe Cys
        195                 200                 205

Gly Ala Gly Val Lys Asn Ala Arg Ala Gln Val Leu Glu Leu Ala Glu
    210                 215                 220

Lys Ile Lys Ser Pro Ile Gly His Ala Leu Gly Gly Lys Gln Tyr Ile
225                 230                 235                 240

Gln His Glu Asn Pro Phe Glu Val Gly Met Ser Gly Leu Leu Gly Tyr
                245                 250                 255

Gly Ala Cys Val Asp Ala Ser Asn Glu Ala Asp Leu Leu Ile Leu Leu
            260                 265                 270

Gly Thr Asp Phe Pro Tyr Ser Asp Phe Leu Pro Lys Asp Asn Val Ala
        275                 280                 285

Gln Val Asp Ile Asn Gly Ala His Ile Gly Arg Arg Thr Thr Val Lys
    290                 295                 300

Tyr Pro Val Thr Gly Asp Val Ala Ala Thr Ile Glu Asn Ile Leu Pro
305                 310                 315                 320

His Val Lys Glu Lys Thr Asp Arg Ser Phe Leu Asp Arg Met Leu Lys
                325                 330                 335

Ala His Glu Arg Lys Leu Ser Ser Val Val Glu Thr Tyr Thr His Asn
            340                 345                 350

Val Glu Lys His Val Pro Ile His Pro Glu Tyr Val Ala Ser Ile Leu
        355                 360                 365

Asn Glu Leu Ala Asp Lys Asp Ala Val Phe Thr Val Asp Thr Gly Met
    370                 375                 380

Cys Asn Val Trp His Ala Arg Tyr Ile Glu Asn Pro Glu Gly Thr Arg
385                 390                 395                 400

Asp Phe Val Gly Ser Phe Arg His Gly Thr Met Ala Asn Ala Leu Pro
                405                 410                 415

His Ala Ile Gly Ala Gln Ser Val Asp Arg Asn Arg Gln Val Ile Ala
            420                 425                 430

Met Cys Gly Asp Gly Gly Leu Gly Met Leu Leu Gly Glu Leu Leu Thr
        435                 440                 445

Val Lys Leu His Gln Leu Pro Leu Lys Ala Val Val Phe Asn Asn Ser
    450                 455                 460

Ser Leu Gly Met Val Lys Leu Glu Met Leu Val Glu Gly Gln Pro Glu
465                 470                 475                 480

Phe Gly Thr Asp His Glu Glu Val Asn Phe Ala Glu Ile Ala Ala Ala
                485                 490                 495
```

Ala Gly Ile Lys Ser Val Arg Ile Thr Asp Pro Lys Lys Val Arg Glu
500 505 510

Gln Leu Ala Glu Ala Leu Ala Tyr Pro Gly Pro Val Leu Ile Asp Ile
515 520 525

Val Thr Asp Pro Asn Ala Leu Ser Ile Pro Pro Thr Ile Thr Trp Glu
530 535 540

Gln Val Met Gly Phe Ser Lys Ala Ala Thr Arg Thr Val Phe Gly Gly
545 550 555 560

Gly Val Gly Ala Met Ile Asp Leu Ala Arg Ser Asn Ile Arg Asn Ile
565 570 575

Pro Thr Pro

<210> 3
<211> 875
<212> DNA
<213> Corynebacterium glutamicum

<400> 3
tgcgagatgg tgaatggtgg tgagcagggt gaacgcattt tgcatcacgc gattcagtcc 60
accatggcgg gtaaaggtgt gtcggtggta gtgattcctg gtgatatcgc taaggaagac 120
gcaggtgacg gtacttattc caattccact atttcttctg gcactcctgt ggtgttcccg 180
gatcctactg aggctgcagc gctggtggag gcgattaaca acgctaagtc tgtcactttg 240
ttctgcggtg cgggcgtgaa gaatgctcgc gcgcaggtgt tggagttggc ggagaagatt 300
aaatcaccga tcgggcatgc gctgggtggt aagcagtaca tccagcatga gaatccgttt 360
gaggtcggca tgtctggcct gcttggttac ggcgcctgcg tggatgcgtc caatgaggcg 420
gatctgctga ttctattggg tacggatttc ccttattctg atttccttcc taaagacaac 480
gttgcccagg tggatatcaa cggtgcgcac attggtcgac gtaccacggt gaagtatccg 540
gtgaccggtg atgttgctgc aacaatcgaa aatattttgc ctcatgtgaa ggaaaaaaca 600
gatcgttcct tccttgatcg gatgctcaag gcacacgagc gtaagttgag ctcggtggta 660
gagacgtaca cacataacgt cgagaagcat gtgcctattc accctgaata cgttgcctct 720
attttgaacg agctggcgga taaggatgcg gtgtttactg tggataccgg catgtgcaat 780
gtgtggcatg cgaggtacat cgagaatccg gagggaacgc gcgactttgt gggttcattc 840
cgccacggca cgatggctaa tgcgttgcct catgc 875